DOCUMENT RESUME

ED 405 380 TM 026 246

AUTHOR Schiel, Jeff
TITLE Student Effort and Performance on a Measure of Postsecondary Educational Development. ACT Research Report Series 96-9.
INSTITUTION American Coll. Testing Program, Iowa City, Iowa.
PUB DATE Dec 96
NOTE 28p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Entrance Examinations; *College Students; Ethnic Groups; Higher Education; High School Graduates; Institutional Characteristics; Racial Differences; Scores; Sex Differences; Standardized Tests; *Student Motivation; *Test Results
IDENTIFIERS ACT Assessment; *Collegiate Assessment of Academic Proficiency; *Effort

ABSTRACT

This study examined the relationship between student effort and performance on the Collegiate Assessment of Academic Proficiency (CAAP), a standardized measure of postsecondary educational development, while statistically controlling for American College Testing (ACT) Assessment scores, ethnicity, gender, length of time between ACT Assessment and CAAP testing, and type of institution attended. Data were analyzed for 50,786 students at 188 postsecondary institutions. Results show that students who reported giving reasonable effort while taking the test earned scores that were, on average, one-half to one and one-fourth standard deviation units higher than those of students who reported giving no effort. This occurred irrespective of students' ACT Assessment scores and other selected independent variables. (Contains four tables, six figures, and four references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ACT Research Report Series 96-9

ED 405 380

Student Effort and Performance on a Measure of Postsecondary Educational Development

Jeff Schiel

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
P. H. FARRANT
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

TM026246

ACT

2 December 1996

For additional copies write:
ACT Research Report Series
PO Box 168
Iowa City, Iowa 52243-0168

© 1997 by ACT, Inc. All rights reserved.

Student Effort and Performance on a Measure of Postsecondary Educational Development

Jeff Schiel

Abstract

This study examined the relationship between student effort and performance on the CAAP, a standardized measure of postsecondary educational development, while statistically controlling for ACT Assessment scores, ethnicity, gender, length of time between ACT Assessment and CAAP testing, and type of institution attended. Data were analyzed for 50,786 students at 188 postsecondary institutions. The results showed that students who reported giving reasonable effort while testing earned scores that were, on average, ½ to 1¼ standard deviation units higher than those of students who reported giving no effort. This occurred irrespective of students' ACT Assessment scores and other selected independent variables.

Student Effort and Performance on a Measure of Postsecondary Educational Development

An integral part of many postsecondary outcomes assessment programs and accreditation efforts is measuring, by way of objective or performance measures, students' educational development in relevant areas. It is essential to the success of these efforts that students be motivated to perform effectively while testing. If students do not give reasonable effort, then it is likely that the results of the testing will be adversely influenced by measurement error and therefore cannot be considered valid.

Different strategies can be used to motivate students to give their best efforts in the testing situation, including awarding scholarships, attaching the resulting test scores to students' academic records, and emphasizing the benefits of assessment. Too often, however, institutions choose ineffective motivating strategies or neglect to consider them altogether. This is particularly true when institutions first engage in outcomes assessment, and typically occurs because staff and administrators are unfamiliar with, or fail to consider, the unfavorable effects of poor student effort.

Research at the secondary level has shown that when students were motivated to do their best on group achievement tests, their scores increased substantially (Taylor & White, 1981). It seems reasonable that a similar relationship between level of effort and test performance would also exist at the postsecondary level, but there is little evidence that this is so. A literature review revealed no published empirical research examining the relationship of student effort and performance on postsecondary outcomes measures. An unpublished study indicated that group mean scores on the College Outcomes Measures Program were noticeably lower when 15% or more of examinees reported that they gave little effort (Steele, 1996). This study, although adjusting to some

extent for entering level of achievement, did not control statistically for other student background characteristics.

The ACT Assessment is a curriculum-based test of educational development that is used for college admissions and placement. The Collegiate Assessment of Academic Proficiency (CAAP) is designed to measure the educational development of college sophomores. It is typically used in outcomes assessment. The purpose of the present study was to examine the relationship of student effort and performance on the CAAP, while statistically controlling for such variables as ACT Assessment scores, ethnicity, gender, length of time between ACT Assessment and CAAP testing, and type of institution (two- or four-year) or specific institution attended.

Data

The sample for the study consisted of 50,786 students representing 188 postsecondary institutions who took the CAAP between August 1992 and June 1996 and who completed the ACT Assessment between September 1990 and September 1994. Although nearly all students in the sample (94%) took the ACT Assessment on national test dates, some students tested under extended-time conditions (an option offered to students with physical disabilities) or were tested soon after they enrolled at a college or university. These students were included in order to more accurately represent the population of students who take both the ACT Assessment and the CAAP.

There are five CAAP objective tests (Writing Skills, Mathematics, Reading, Critical Thinking, and Science Reasoning) and a Writing Essay test. CAAP scores range from 40 to 80, with a mean of approximately 60 and a standard deviation of approximately

5. CAAP is a modular testing program; one or more of the tests may be administered, depending on an institution's needs and resources. Because of this characteristic, separate analyses were conducted on each of the CAAP objective tests. The number of students for each test ranged from 20,420 (CAAP Critical Thinking) to 33,543 (CAAP Mathematics). The number of institutions represented by these students ranged from 101 (CAAP Science Reasoning) to 152 (CAAP Mathematics). Minimum sample size was set at 25 students per institution.

The exact date of CAAP testing is not recorded in the CAAP data file. However, a scoring date is recorded; this usually occurs within one to four weeks after testing. The CAAP scoring date and the ACT Assessment test date were used to determine the approximate length of time (in months) between ACT Assessment and CAAP testing.

Institutions interested in measuring educational change over time sometimes test incoming freshmen with the CAAP and then retest them at the end of their sophomore year when they have completed their general education course work. Other institutions do not test incoming freshmen at all, but focus exclusively on students near the end of their sophomore year or students who have completed a specific core of courses or number of credit hours. Most of the students in this study (about 57%) reported that they took CAAP during their sophomore year, but some reported that they took it when they were freshmen or juniors (19% and 21%, respectively). About 3% took CAAP during their senior year. Because postsecondary institutions may administer CAAP to students who are at different educational levels, and because the ACT Assessment may be taken during either the junior or senior year of high school, the length of time

between ACT Assessment and CAAP testing varies. The sample of students was therefore limited to those who took the CAAP at least one month after taking the ACT Assessment. Length of time between ACT Assessment and CAAP testing was taken into account in the analyses.

The matched ACT Assessment/CAAP data files contained ACT Assessment and CAAP test scores, the CAAP scoring date, the ACT Assessment test date, student background information (i.e., gender, ethnicity), institutional type (two-year, four-year), institution attended, and responses to a question about the level of students' efforts at the time of CAAP testing. After students completed the CAAP, they chose one of four possible responses to describe their level of effort: "tried my best," "gave moderate effort," "gave little effort," or "gave no effort." For all students who took both the ACT Assessment and the CAAP during 1990-96 (n = 71,416), the response rate for this question was about 88%. For the primary analyses, responses to this question were recoded to "gave no effort" (0) or "any effort at all" (any of the first three responses = 1). Although a small amount of information was lost by coding the student effort variable in this fashion, a considerable degree of interpretability was gained. However, for comparison purposes, alternative codings of the student effort variable were investigated: coding it as 0, 1, 2, or 3 or using different dummy codings (e.g., gave no effort or gave little effort = 0, gave moderate effort or tried my best = 1).

Ethnicity, gender, and institutional type were also dummy coded. Ethnicity was coded so that the regression coefficients for the individual ethnic groups (African American, Asian American, Hispanic or Native American) could be interpreted as the

difference in average CAAP score between each ethnic group and Caucasian American students. Gender was coded as females = 1 and males = 0. Institutional type was coded as two-year = 0 and four-year = 1.

Method

Descriptive statistics were calculated for ACT Assessment and CAAP scores and other relevant variables. Correlation coefficients between CAAP scores and relevant independent variables were also calculated.

Frequencies of the student effort variable responses were computed, by institution. These would help to determine, for example, whether students giving no effort on the CAAP were evenly distributed across all institutions, or whether they were from only a few institutions.

Preliminary regression models were developed by regressing CAAP scores on ACT Assessment scores, ethnicity, gender, institutional type, length of time (in months) between ACT Assessment and CAAP testing, and student effort. The regression models were evaluated in terms of model statistical significance ($p < .001$), collinearity of the independent variables, and the statistical significance ($p < .001$) of the regression coefficient associated with each independent variable. Using multiple regression, the regression coefficient associated with the dummy-coded student effort variable can be interpreted as reflecting the difference in average CAAP score associated with giving some effort on the test.

The principal goal in developing regression models was to explain the effect of student effort on CAAP scores while statistically controlling for other relevant

independent variables. There was relatively less interest in finding the most accurate models for estimating CAAP scores. As a result, relatively few independent variables were investigated and subsequently used in the regression models.

Institutional Type/Institution Attended

One might expect the average level of student effort within two different institutions to differ somewhat, even if similar motivating strategies were employed. For example, student effort at two-year institutions might differ from that at four-year institutions. Institutional type and institution attended are therefore important variables to include when modeling CAAP scores. Institution attended, because it would more directly reflect students' educational experiences, would likely provide relatively more useful information for modeling CAAP performance. In comparison, institutional type would be less precise because it consists of only two categories. However, it would yield results that would be relatively easy to interpret (e.g., two regression coefficients, representing two- and four-year institutions vs. 100 or more coefficients representing specific institutions). Therefore, separate analyses for modeling CAAP scores were conducted using either institutional type or institution attended.

Because ACT Assessment scores were to be included as covariates (to control for preexisting differences in students' educational development), including institutional type or institution attended in the models was contingent upon satisfying the assumption of homogeneous slopes across institutions. If heterogeneous slopes were found, then this would necessitate fitting separate regression equations for two- and four-year institutions or for each institution.

To test the assumption of homogeneous slopes, institutional type/institution attended by ACT Assessment score interaction terms were included in models containing ACT Assessment score and institutional type/institution attended main effects. Separate models were developed for each CAAP test. Most institutional type by ACT Assessment score interaction terms for these models were not statistically significant ($p \geq .001$); as a result, separate within-institutional type regression equations were not developed for the models based on institutional type.

In comparison, all institution attended by ACT Assessment score interactions were statistically significant. However, because of the very large sample sizes, statistical significance may not be a meaningful indicator of heterogeneous slopes. To assist in determining whether slopes were heterogeneous, the range of the regression coefficients for the interaction terms was examined for each model. Ranges were fairly large (e.g., > 0.9 CAAP score units) for four of the five models, suggesting that the slopes differed meaningfully across institutions. A separate analysis was therefore done in which regression equations were fitted for each institution and the results summarized across institutions. The regression equations for each institution used the same independent variables described below for the final models, except that institutional type was not included.

Final Model Development

All independent variables except student effort were entered first into the regression equations. These variables illustrated conditions that could not be changed or manipulated at the time of CAAP testing. Students would have some control over

the amount of effort they gave while testing, and so this variable was entered last into the regression equations. The final models were:

CAAP Writing Skills	=	*f* (ACT English score, ethnicity, gender, length of time between ACT Assessment and CAAP testing, institutional type, student effort)
CAAP Mathematics	=	*f* (ACT Mathematics score, ethnicity, gender, student effort)
CAAP Reading	=	*f* (ACT Composite score, ethnicity, gender, length of time between ACT Assessment and CAAP testing, student effort)
CAAP Critical Thinking	=	*f* (ACT Composite score, ethnicity, gender, length of time between ACT Assessment and CAAP testing, student effort)
CAAP Science Reasoning	=	*f* (ACT Composite score, ethnicity, gender, length of time between ACT Assessment and CAAP testing, institutional type, student effort).

Results

Descriptive Statistics

Means and percentages for the variables included in the regression models are shown in Table 1. Statistics are not reported if the variables were not included in final regression models.

TABLE 1

Descriptive Statistics for Variables in CAAP Test Score Models

		CAAP				
Statistic	Variable	Writing Skills (n=33,222)	Mathematics (n=33,543)	Reading (n=29,443)	Critical Thinking (n=20,420)	Science Reasoning (n=21,001)
Mean	CAAP score	63.9	57.5	62.1	61.7	59.7
	ACT Assessment score[1]	20.9	20.0	21.2	21.0	21.3
	Months between ACT & CAAP testing	31.3	--	30.8	28.1	30.4
Std. dev.	CAAP score	4.6	3.8	5.3	5.3	4.4
	ACT Assessment score	4.8	4.4	4.3	4.3	4.3
	Months between ACT & CAAP testing	11.2	--	11.3	12.2	11.2
Percent	Ethnicity					
	African American	13	13	14	12	11
	Asian American	1	1	1	1	1
	Caucasian American	83	83	81	81	85
	Hispanic, Native American	3	3	4	6	3
	Gender					
	Female	62	61	62	63	61
	Male	38	39	38	37	40
	Institutional type					
	Two-year	32	--	--	--	27
	Four-year	69	--	--	--	73
	Student effort					
	Tried my best	69	45	47	51	32
	Gave moderate effort	27	39	39	38	41
	Gave little effort	3	14	11	9	22
	Gave no effort	1	3	3	2	6

[1]ACT English and Mathematics scores were used to model CAAP Writing Skills and Mathematics scores, respectively. ACT Composite score was used to model CAAP Reading, Science Reasoning, and Critical Thinking scores.

Mean CAAP scores for the students in this study were comparable to those reported in the CAAP user norms (ACT, 1995a). Mean ACT Composite scores were comparable to those reported for enrolled college freshmen (ACT, 1995b). Percentages of ethnic and gender groups were fairly similar across CAAP test scores.

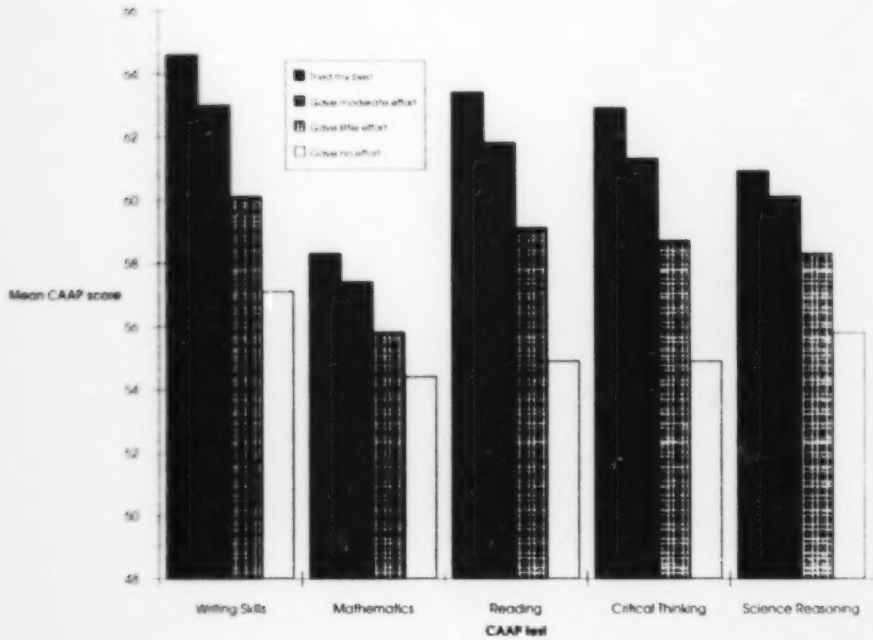
BEST COPY AVAILABLE

Fairly small percentages of students reported that they gave no effort on the CAAP tests; these percentages ranged from 1% (Writing Skills) to 6% (Science Reasoning). Percentages of students indicating that they gave no effort were fairly evenly distributed across institutions.

Institutions administering more than one CAAP test were asked, as part of the test administration procedure, to administer the tests in the following order: Writing Skills, Mathematics, Reading, Critical Thinking, Science Reasoning. Because the Science Reasoning test is always taken last when two or more tests are administered, it is possible that student effort on this test might, at times, be relatively low due to factors such as fatigue. This could account for the relatively larger percentage of students reporting that they gave no effort on the Science Reasoning test.

Figure 1 shows mean CAAP scores, by level of student effort. For all tests except Mathematics, the largest difference in mean scores between any two adjacent response categories occurred for "gave little effort" and "gave no effort." For the Reading test, for example, the difference in mean scores between these two categories was 4.2 scale score units. The difference in mean scores between "gave moderate effort" and "gave little effort" was slightly smaller (2.7). The relatively large difference in mean scores between the categories of "gave little effort" and "gave no effort" provides support for coding the student effort variable in the primary analysis as "gave no effort" (0) or "any effort at all" (1).

FIGURE 1. Mean CAAP Scores, by Level of Student Effort



Correlation Coefficients

Point-biserial correlations between the student effort dummy variable and CAAP scores ranged from .13 (Writing Skills, Mathematics) to .22 (Reading, Science Reasoning). When student effort was coded as 0, 1, 2, or 3, Pearson product-moment correlations between this variable and CAAP scores were slightly larger, ranging from .24 (Writing Skills) to .33 (Reading).

Regression Analyses

Regression coefficients for all independent variables are summarized in Table 2, along with multiple R and standard error of estimate (SEE) for each CAAP model. Multiple R was fairly similar across models, ranging from .73 (Science Reasoning) to .79

(Writing Skills). SEE was smallest for the Mathematics model (2.53) and largest for the Critical Thinking model (3.57).

TABLE 2

Regression Statistics for CAAP Test Score Models

	CAAP				
Statistic	**Writing Skills**	**Mathematics**	**Reading**	**Critical Thinking**	**Science Reasoning**
R	.79	.74	.75	.74	.73
SEE	2.81	2.53	3.55	3.57	3.03
Regression coefficients					
Intercept	42.23	43.41	35.03	35.58	40.90
ACT score[1]	.70	.61	.88	.87	.70
Ethnicity					
African American	-1.46	-.04*	-.62	-.42	-.80
Asian American	-.72	.33*	-1.02	-1.48	.32*
Hispanic, Native American	-.69	-.06*	-.05*	-.32*	.00*
Gender	.57	-.23	.68	.30	-.87
Months between ACT & CAAP testing	.02	--	.05	.05	.03
Institutional type	.17	--	--	--	-.19
Student effort	6.19	1.99	6.47	6.45	3.81

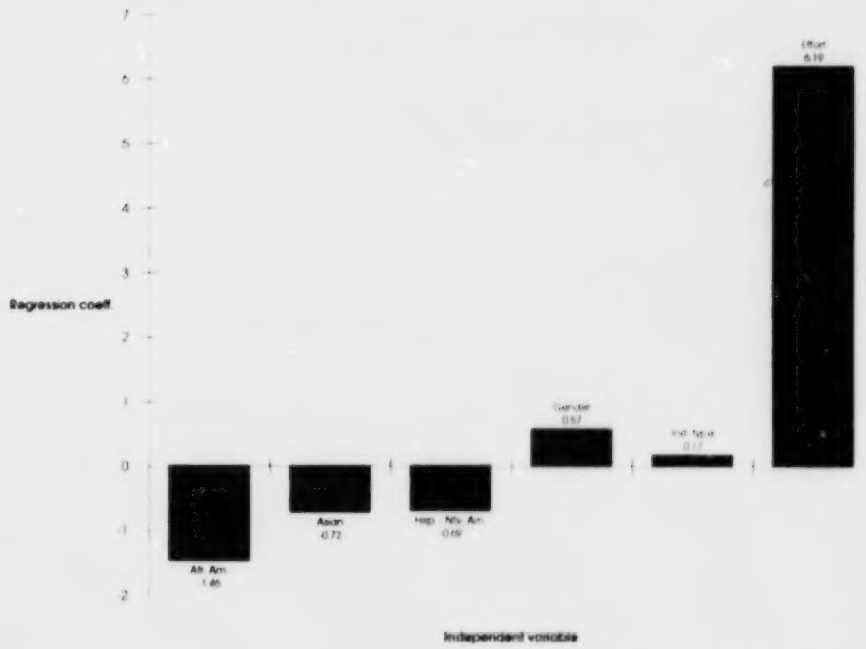
Notes: [1]ACT English and Mathematics scores were used to model CAAP Writing Skills and Mathematics scores, respectively. ACT Composite score was used to model CAAP Reading, Science Reasoning, and Critical Thinking scores. *Not statistically significant ($p \geq .001$).

Not all of the regression coefficients for ethnicity were statistically significant for all models. Because the coefficients for each individual ethnic group were part of a system for dummy coding the ethnicity variable, all coefficients were left in the models for illustrative purposes.

Figures 2 through 6 illustrate the relative size of the regression coefficients for the dummy-coded independent variables in the CAAP test score models. Given all other variables in the model, the regression coefficient associated with the student effort

variable for the CAAP Writing Skills test was very large (6.19), relative to those associated with other dummy-coded variables in the model (see Figure 2). This coefficient indicated that students who reported giving at least some effort on the Writing Skills test could expect, on average, to score about six scale score units higher than students who reported giving no effort. This was true regardless of their ACT Assessment scores, ethnicity, gender, length of time between ACT Assessment and CAAP testing, and type of institution attended.

FIGURE 2. Regression Coefficients for Dummy-Coded Independent Variables in CAAP Writing Skills Model



The designation "ns" in Figures 3 through 6 indicates that a particular ethnicity regression coefficient was not statistically significant ($p \geq .001$). Institutional type was not statistically significant for the CAAP Mathematics, Reading, and Science Reasoning models (Figures 3-5) and is therefore not shown.

FIGURE 3. Regression Coefficients for Dummy-Coded Independent Variables in CAAP Mathematics Model

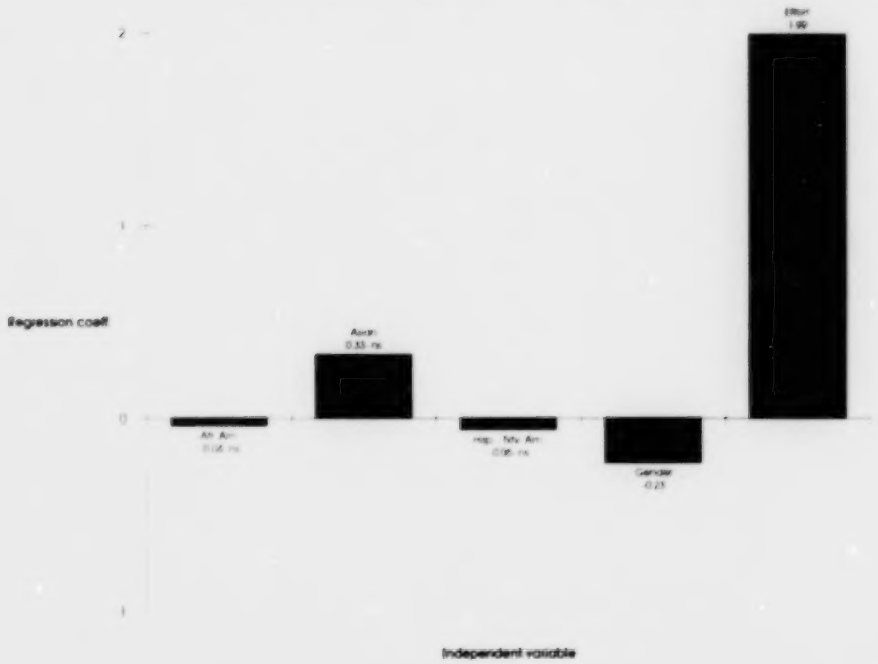


FIGURE 4. Regression Coefficients for Dummy-Coded Independent Variables in CAAP Reading Model

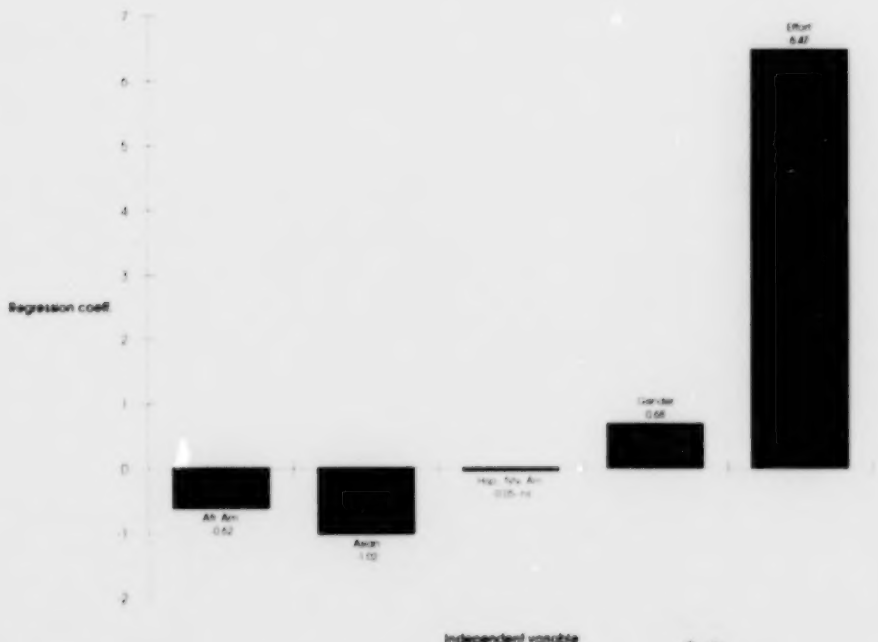


FIGURE 5. Regression Coefficients for Dummy-Coded Independent Variables in CAAP Critical Thinking Model

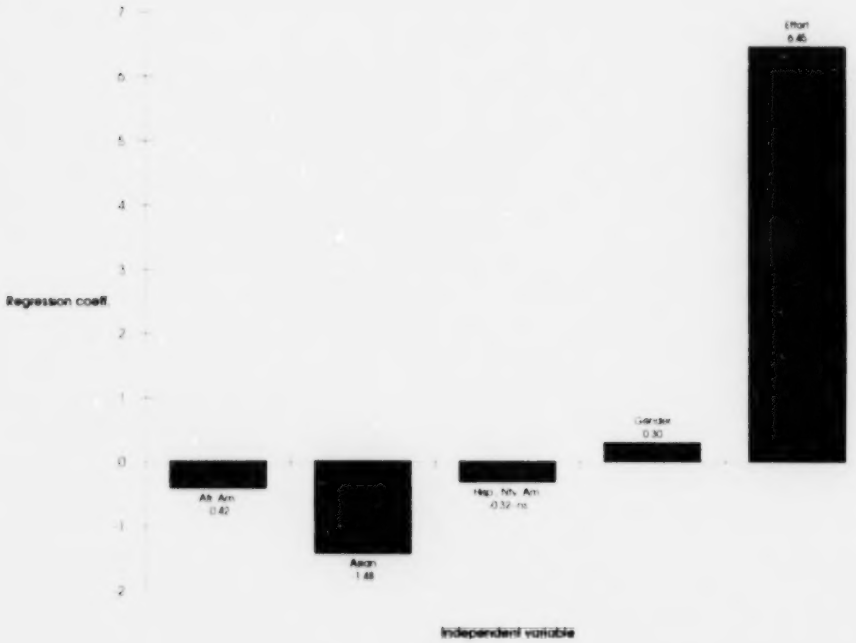
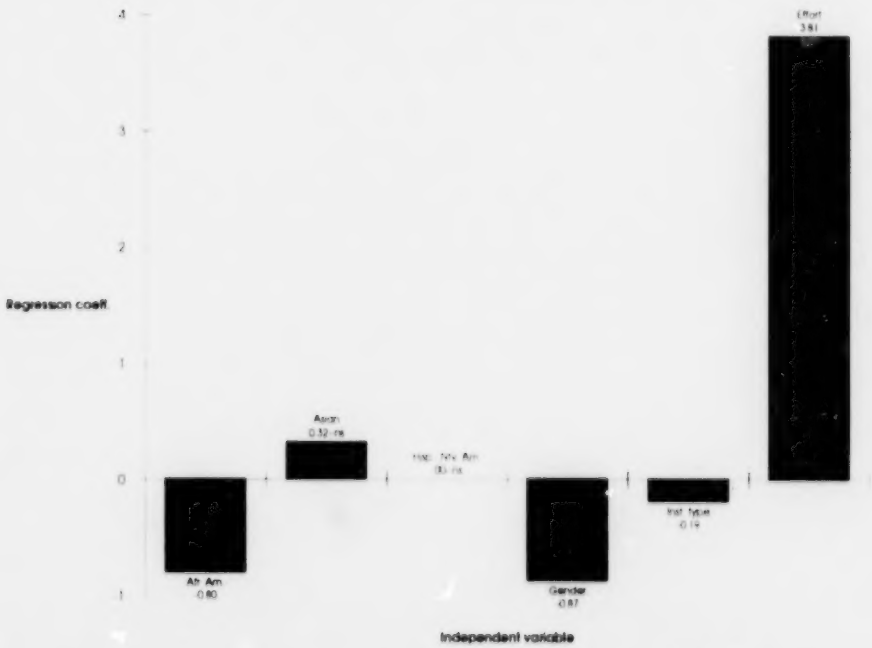


FIGURE 6. Regression Coefficients for Dummy-Coded Independent Variables in CAAP Science Reasoning Model



Figures 3 through 6 illustrate that the regression coefficients associated with student effort for the CAAP Mathematics, Reading, Critical Thinking, and Science Reasoning models (1.99, 6.47, 6.45, and 3.81, respectively) were also large relative to those of other dummy-coded independent variables. Expressed in standard deviation units, differences in average CAAP scores reflected by these coefficients ranged from about ½ to 1¼ CAAP standard deviation units.

Regression coefficients for ethnicity and gender were considerably smaller than those for student effort, indicating that these variables contributed relatively little to CAAP performance, when statistically controlling for other independent variables in the model. Coefficients for ethnicity ranged from -1.48 (Asian Americans, Critical Thinking test) to .33 (Asian Americans, Mathematics test). Caucasian American students typically had higher CAAP scores than did ethnic minority students when all independent variables were statistically controlled. Coefficients for gender ranged from -.87 (Science Reasoning) to .68 (Reading). Females performed somewhat better than males on the CAAP Writing Skills, Reading, and Critical Thinking tests when all independent variables were statistically controlled.

Alternative codings of the student effort variable. Alternative codings of the student effort variable yielded relatively smaller regression coefficients than those of the original coding. For example, coding this variable as "gave little or no effort" = 0 and "gave moderate effort or tried my best" = 1 yielded regression coefficients ranging from 1.22 (Mathematics) to 3.48 (Reading).

Table 3 contains regression coefficients, multiple R, and SEE for CAAP test score models when the student effort variable was coded as 0, 1, 2, or 3. Multiple R based on this coding was, for some models, slightly larger than that based on the original coding (see Table 2). SEE was slightly smaller for all models based on the alternative coding. These results reflect a slight increase in prediction accuracy due to the alternative coding of the student effort variable.

TABLE 3

Regression Statistics for CAAP Test Score Models When the Student Effort Variable Was Coded As 0, 1, 2, or 3

	CAAP				
Statistic	**Writing Skills**	**Mathematics**	**Reading**	**Critical Thinking**	**Science Reasoning**
R	.79	.75	.76	.76	.75
SEE	2.78	2.49	3.44	3.50	2.95
Regression coefficients					
Intercept	45.87	44.22	38.16	38.72	42.36
ACT score[1]	.69	.60	.86	.85	.69
Ethnicity					
African American	-1.49	-.15	-.75	-.52	-.90
Asian American	-.69	.25*	.99	-1.41	.23*
Hispanic, Native American	-.63	-.06*	-.07*	-.29*	.01*
Gender	.37	-.31	.38	.08*	-.92
Months between ACT & CAAP testing	.02	--	.00	.05	.03
Institutional type	.12	--	--	--	-.18
Student effort	1.13	.68	1.80	1.63	1.30

Notes: [1]ACT English and Mathematics scores were used to model CAAP Writing Skills and Mathematics scores, respectively. ACT Composite score was used to model CAAP Reading, Science Reasoning, and Critical Thinking scores. *Not statistically significant ($p \geq .001$).

The regression coefficients for student effort in Table 3 reflect the difference in average CAAP score associated with a one-unit increase in the student effort variable. It is relatively more difficult to interpret a one-unit increase in student effort when this

variable is coded as 0, 1, 2, or 3. The original coding provides a more straightforward interpretation of the regression coefficient associated with student effort.

As described previously, the largest difference in mean CAAP scores between any two adjacent response categories of the student effort variable typically occurred for "gave little effort" and "gave no effort." This likely accounts for the large regression coefficients for the original coding, relative to those of alternative codings.

Controlling for all other independent variables slightly reduced the effect of student effort on CAAP scores, regardless of how the student effort variable was coded. When effort was the only independent variable in the model, regression coefficients for this variable were somewhat larger, ranging from 3.18 (Mathematics) to 7.38 (Reading).

Institution attended. Fitting separate regression equations for each institution yielded median (across institutions) regression coefficients for the student effort variable that were slightly smaller than those found in the models based on institutional type. They ranged from 1.61 (Mathematics) to 5.57 (Reading). The student effort regression coefficients for both analyses are shown in Table 4. These coefficients indicate that the relationship between student effort and CAAP performance was substantial in both analyses.

TABLE 4

Regression Coefficients for Student Effort: Models Based on Institutional Type Vs. Within-Institution Analyses

Analysis	Regression coefficients				
	Writing Skills	Mathematics	Reading	Critical Thinking	Science Reasoning
Within-institution: Separate regression equations fitted for each institution; median (min./max.) across institutions	5.10 (-3.76/13.00)	1.61 (-3.56/6.09)	5.57 (-.67/8.68)	5.56 (-1.51/15.03)	2.95 (-.30/8.43)
Institutional type (two-year, four-year) plus all other independent variables included in one regression equation	6.19	1.99	6.47	6.45	3.81

The minimum regression coefficients in Table 4 indicate that the within-institution analysis yielded, for some institutions, negative regression coefficients for the student effort variable. The percentage of institutions with negative regression coefficients for this variable ranged from about 1% (Science Reasoning) to 3% (Mathematics). One possible explanation for these negative coefficients is that students at these institutions were not at all motivated to give reasonable effort while taking the CAAP and did not take the testing situation seriously (e.g., they actually gave no effort and received low CAAP scores, but falsely reported that they did give some effort). Regardless of the reasons for the negative coefficients occurring, their effect on the results is likely minimal, given the large number of institutions in this study.

Outlier Analysis

The student effort data in this study were self-reported. It is likely that some students who reported giving no effort actually gave some effort and vice versa. For example, two students who reported giving no effort earned CAAP Writing Skills scores

of 71; this particular score is equal to or higher than the scores of 97% of all CAAP-tested four-year college sophomores.

To further examine the relationship of student effort and CAAP score, an outlier analysis was performed, by institution. This analysis revealed the presence of only a few influential outliers. For example, 15 outliers were found in the Writing Skills data file (n = 33,222). These were distributed across multiple institutions. Although it was possible that the removal of outliers would have slightly improved the fit of regression models, there was little justification for so doing. The goal of this study was to explain the effect of student effort on test performance, rather than predicting test performance from student effort as accurately as possible. In addition, there was no way to determine whether the outliers were coding errors or accurate representations of students' opinions of their levels of effort. For these reasons, the outliers were not removed.

Discussion

The results of this study emphasize the importance of motivating students to give reasonable effort (i.e., give at least little or moderate effort, or try their best) while taking tests like the CAAP. Students who gave reasonable effort while testing earned, on average, considerably higher CAAP scores (about ½ to 1¼ CAAP standard deviation units higher) than did students who gave no effort. This occurred irrespective of students' ACT Assessment scores, ethnicity, gender, length of time between ACT Assessment and CAAP testing, and type of institution attended.

Statistically controlling for institution attended slightly reduced the effect of student effort on CAAP score, relative to the effect found when institutional type (i.e., two-year or four-year) was statistically controlled. Although fitting within-school regression equations controlled, to some degree, for differences in institutional motivating strategies, no information was available about the specific strategies used by institutions. Future research in this area could benefit from the collection of information about institutional motivating strategies.

The student effort data were based on students' self-reports of their levels of effort at the time of CAAP testing. The outlier analysis detected only a few influential observations in which a relatively high CAAP score was associated with giving no effort on the test, or vice versa. Such an analysis does not, of course, completely address the question of whether student effort data were accurately reported. Because the extent to which students accurately reported their level of effort is not fully known, the results of this study must be interpreted with this limitation in mind.

Implications

The results of this study suggest that, particularly in the context of outcomes assessment or accreditation efforts, testing students who are not sufficiently motivated to give reasonable effort could yield anomalous and invalid results. For example, consider an institution whose staff have performed careful content evaluations of its required mathematics courses and a standardized mathematics test. Given the results of the evaluations, the test appears capable of measuring the mathematical skills and knowledge that students should obtain as a result of completing the mathematics

requirements. The staff therefore decide to use the standardized test, but fail to motivate students to give reasonable effort while testing. After testing all of their students and ascertaining that each has, in fact, completed the mathematics requirements, they discover that their institution's average mathematics test score is considerably lower than those of comparable institutions. This puzzling result might lead staff to conclude, perhaps erroneously, that their required mathematics courses are somehow less rigorous than those of other institutions, or that the test is not actually measuring their students' mathematical skills and knowledge. This could have a direct effect on recommendations for curriculum revision. Thoughtful planning and implementation of effective student motivating strategies would increase the likelihood of achieving valid results.

References

American College Testing (1995a). CAAP user norms, Fall 1995. Iowa City, IA: Author.

American College Testing (1995b). The ACT Class Profile National Norms. Iowa City, IA: Author.

Steele, J. M. (1996). How do you get 'um to try? Who--students? No--faculty and administrators! A paper presented at the annual meeting of the American Evaluation Association.

Taylor, C. and White, K. R. (1981). Effects of reinforcement and training on Title I students' group standardized test performance. A paper presented at the annual meeting of the American Educational Research Association.